



# Modeling the shape hierarchy for visually guided grasping

Omid Rezaei<sup>1\*</sup>, Ashley Kleinhans<sup>2,3</sup>, Eduardo Matallanas<sup>4</sup>, Ben Selby<sup>1</sup> and Bryan P. Tripp<sup>1\*</sup>

<sup>1</sup> Department of Systems Design Engineering, Centre for Theoretical Neuroscience, University of Waterloo, Waterloo, ON, Canada

<sup>2</sup> Mobile Intelligent Autonomous Systems, Council for Scientific and Industrial Research, Pretoria, South Africa

<sup>3</sup> School of Mechanical and Industrial Engineering, University of Johannesburg, Johannesburg, South Africa

<sup>4</sup> ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain

## Edited by:

Antonio J. Rodriguez-Sanchez,  
University of Innsbruck, Austria

## Reviewed by:

Abdelmalik Moujahid, University of  
the Basque Country UPV/EHU,  
Spain

Ko Sakai, University of Tsukuba,  
Japan

Peter Janssen, Catholic University  
of Leuven, Belgium

## \*Correspondence:

Omid Rezaei and Bryan P. Tripp,  
Systems Design Engineering,  
University of Waterloo, 200  
University Avenue West, Waterloo,  
ON N2L 3G1, Canada  
e-mail: [omid.srezaei@uwaterloo.ca](mailto:omid.srezaei@uwaterloo.ca);  
[bptrip@uwaterloo.ca](mailto:bptrip@uwaterloo.ca)

The monkey anterior intraparietal area (AIP) encodes visual information about three-dimensional object shape that is used to shape the hand for grasping. We modeled shape tuning in visual AIP neurons and its relationship with curvature and gradient information from the caudal intraparietal area (CIP). The main goal was to gain insight into the kinds of shape parameterizations that can account for AIP tuning and that are consistent with both the inputs to AIP and the role of AIP in grasping. We first experimented with superquadric shape parameters. We considered superquadrics because they occupy a role in robotics that is similar to AIP, in that superquadric fits are derived from visual input and used for grasp planning. We also experimented with an alternative shape parameterization that was based on an Isomap dimension reduction of spatial derivatives of depth (i.e., distance from the observer to the object surface). We considered an Isomap-based model because its parameters lacked discontinuities between similar shapes. When we matched the dimension of the Isomap to the number of superquadric parameters, the superquadric model fit the AIP data somewhat more closely. However, higher-dimensional Isomaps provided excellent fits. Also, we found that the Isomap parameters could be approximated much more accurately than superquadric parameters by feedforward neural networks with CIP-like inputs. We conclude that Isomaps, or perhaps alternative dimension reductions of visual inputs to AIP, provide a promising model of AIP electrophysiology data. Further work is needed to test whether such shape parameterizations actually provide an effective basis for grasp control.

**Keywords:** AIP, CIP, grasping, 3D shape, cosine tuning, superquadrics, Isomap

## 1. INTRODUCTION

The macaque anterior intraparietal area (AIP) receives input from the visual cortex, and is involved in visually guided grasping. A large fraction of neurons in this area encode information about three-dimensional object shapes from visual input (Murata et al., 2000; Sakaguchi et al., 2010). Responses are typically relatively invariant to object position in depth (Srivastava et al., 2009). The responses of some neurons are also invariant to other properties. For example, some are orientation-tuned but not highly sensitive to object shape (Murata et al., 2000). AIP has a strong recurrent connection with premotor area F5, which is involved in hand shaping for grasping (Rizzolatti et al., 1990; Luppino et al., 1999; Borra et al., 2008). Reversible inactivation of AIP leads to grasping impairment, specifically a mismatch between object shape and hand preshape (Gallese et al., 1994; Fogassi et al., 2001). AIP is therefore thought to provide visual information for grasp control (Jeannerod et al., 1995; Fagg and Arbib, 1998).

The focus of this paper is the pathway from V3 and V3A, to the caudal intraparietal area (CIP), to visual-dominant neurons in AIP (Nakamura et al., 2001; Tsutsui et al., 2002). This pathway makes binocular disparity information available for grasp control. Most V3 neurons are selective for binocular disparity (Adams

and Zeki, 2001). V3 sends a major projection to V3A (Felleman et al., 1997), which is also strongly activated during binocular disparity processing (Tsao et al., 2003). Both V3 and V3A project to CIP (Katsuyama et al., 2010). CIP neurons are selective for depth gradients (Taira et al., 2000; Tsutsui et al., 2002; Rosenberg et al., 2013) and curvature (Katsuyama et al., 2010). Neurons in AIP receive disynaptic input from V3A via CIP (Nakamura et al., 2001; Borra et al., 2008). Visual-dominant AIP neurons are selective for 3D object shape (Srivastava et al., 2009; Sakaguchi et al., 2010) cued by binocular disparity, consistent with input from this pathway.

AIP also receives many other inputs that we do not model in the present study. The first of these is the premotor area F5, which together with AIP forms a circuit for grasp-related visuomotor transformations. AIP also receives input from the second somatosensory (SII) cortical region (Krubitzer et al., 1995; Fitzgerald et al., 2004; Gregoriou et al., 2006), which may provide tactile feedback and memory-based somatosensory expectations for grasping. Strong connections with other parietal areas are also identified, as well as with prefrontal areas 46 and 12. Area 12 is implicated in high level non-spatial processing including encoding of objects in working memory, suggesting that AIP may be influenced by visual memory of object features

(Borra et al., 2008). AIP also contains other neurons that fire in conjunction with motor plans in addition to or instead of visual input (Sakata et al., 1997; Murata et al., 2000; Taira et al., 2000). Interestingly, AIP also receives subcortical input (via the thalamus) from both the cerebellum and basal ganglia (Clower et al., 2005). Finally, AIP receives input from the inferotemporal cortex (IT), which is likely to provide additional visual information about shapes. Our present focus however is the visual input from CIP.

The main goal of this study is to model the neural spike code of object-selective visual-dominant AIP neurons. In particular, we wanted to know whether there are certain sets of shape parameters that are consistent with the responses of visual AIP neurons, and which can furthermore be estimated in a physiologically plausible way from the information available in CIP.

We therefore compared two ways of parameterizing shapes. First we considered the superquadric family of shapes, a continuum that includes cuboids, ellipsoids, spheres, octahedra, and cylinders, and which can also be extended in various ways to model more complex shapes (Solina and Bajcsy, 1990). We considered superquadrics because they play a role in robotic grasp control (Duncan et al., 2013) that seems to be similar to the role of AIP in primate grasp control, i.e., they represent shapes compactly as a basis for grasp planning. We also considered an alternative shape parameterization that is based on non-linear dimension reduction of the depth field. In particular, we used an Isomap (Tenenbaum et al., 2000). We considered Isomap parameters partly because they are continuous, i.e., similar shapes have similar parameters. This is consistent with datasets in which similar 3D stimuli elicit similar spike rate patterns in AIP (Theys et al., 2012, Figure 10; Srivastava et al., 2009, Figure 11C).

This study is one of the first to model the mapping from CIP to AIP. Oztop et al. (2006) modeled AIP as a hidden layer in a multi-layer perceptron network that mapped visual depth onto hand configuration. The output layer of this model (corresponding to F5) was a self-organizing map of subnetworks that corresponded to different hand configurations. Prevete et al. (2011) developed a mixed neural and Gaussian-mixture model in which AIP received monocular infero-temporal input. This model did not include stereoscopic input from CIP. The FARS grasping model (Fagg and Arbib, 1998) did not address in detail how AIP activity arises from visual input. While past AIP models have been relatively abstract, here our goal is to fit published tuning curves from AIP recordings, and furthermore to do so using depth-related input from a model of CIP. As far as we are aware, there have not been previous attempts to model AIP tuning in terms of either superquadric parameters or non-linear dimension reduction of depth features.

## 2. MATERIALS AND METHODS

This study consists of three main parts. The first is a model of tuning for depth features in the caudal intraparietal area (CIP, see Section 2.1.1). The second is a model of tuning for three-dimensional shape features in the anterior intraparietal area (AIP, see Section 2.1.2). Finally, the third is an investigation of physiologically plausible feedforward mappings between CIP and AIP (see Section 2.5).

### 2.1. COSINE-TUNING MODELS OF NEUROPHYSIOLOGICAL DATA

We tested how well various tuning curves from the CIP and AIP electrophysiology literature could be approximated by cosine-tuned neuron models. In particular, given a vector  $x$  of stimulus variables, we modeled the net current,  $I$ , driving spiking activity in each neuron as

$$I = \tilde{\phi}^T x + b, \quad (1)$$

where  $b$  is a bias term and  $\tilde{\phi}$  is parallel to the neuron's preferred direction in the space of stimulus parameters. Longer  $\tilde{\phi}$  corresponds to higher sensitivity of the neuron to variations along its preferred direction.

We used a normalized version of the leaky-integrate-and-fire (LIF) spiking model. In this model, the membrane potential  $V$  has subthreshold dynamics  $\tau_{RC} \dot{V} = -V + I$ , where  $\tau_{RC}$  is the membrane time constant and  $I$  is the driving current. The neuron spikes when  $V \geq 1$ , after which  $V$  is held at 0 for a post-spike refractory time  $\tau_{ref}$  before subthreshold integration begins again. These neurons have spike rate

$$r = \frac{1}{\tau_{ref} - \tau_{RC} \cdot \ln(1 - \frac{1}{I})}. \quad (2)$$

Except where noted,  $\tau_{RC}$  was included among the optimization parameters and constrained to the range [0.02s, 0.2s]. In some cases (where noted), when the basic cosine-LIF model (above) produced poor fits, we also added Gaussian background noise to  $I$ . Such background noise more realistically reflects the input to neurons *in vivo* (Carandini, 2004) and causes the LIF model to emit more realistic, irregular spike trains. It also has the potential to produce better tuning curve fits. The reason is that depending on the amplitude of the noise, the spike-rate function may be compressive [as in Equation (2)], sigmoidal, or nearly linear. In these cases we fixed  $\tau_{ref} = 0.005s$  and  $\tau_{RC} = 0.02s$ , included the noise variance as an optimization parameter, and interpolated the spike rate from a lookup table based on simulations. Given a tuning curve from the electrophysiology literature and a list of hypothesized tuning variables, we found least-squares optimal parameters  $\tilde{\phi}$  and  $b$  mainly, and either  $\tau_{RC}$  or  $\sigma_{noise}$  (as noted in the corresponding sections), using Matlab's *lsqcurvefit* function. This function uses Matlab's trust-region-reflective algorithm, which is based partly on Coleman and Li (1994), to solve a non-linear curve-fitting problem in the sense of least-squares. We retried each optimization with at least 1000 random initial points in order to increase the probability of finding a global optimum.

We preferred cosine tuning models over more complex non-linear models for a number of reasons, including that they are simple and that cosine tuning is widespread in the cortex and elsewhere (Zhang and Sejnowski, 1999). (See more detailed rationale in the Discussion).

#### 2.1.1. CIP Tuning

We approximated CIP responses in terms of depth and its first and second spatial derivatives. CIP has been proposed to encode these variables (Orban et al., 2006), and they have been the basis for several experimental studies of CIP responses (Sakata et al.,

1998; Taira et al., 2000; Tsutsui et al., 2001; Katsuyama et al., 2010; Rosenberg et al., 2013).

We fit cosine-tuned LIF neuron models to tuning curves from Tsutsui et al. (2002) and Rosenberg et al. (2013), and from Katsuyama et al. (2010), in which the stimuli varied in terms of first and second derivatives of depth, respectively. The stimuli in Katsuyama et al. (2010) consisted of curved surfaces with depth

$$z = \frac{1}{2} (K_1 x^2 + K_2 y^2). \quad (3)$$

$K_1$  and  $K_2$  were varied to produce two levels of “curvedness,”

$$C = \sqrt{\frac{K_{max}^2 + K_{min}^2}{2}}$$

and a range of “shape indices”

$$SI = \frac{2}{\pi} \arctan \frac{K_{max} + K_{min}}{K_{max} - K_{min}},$$

where  $K_{max}$  and  $K_{min}$  are the larger and smaller curvatures along the  $x$  and  $y$  axes, respectively.

In terms of the depth  $z$ , the principal curvature along the  $x$  axis is

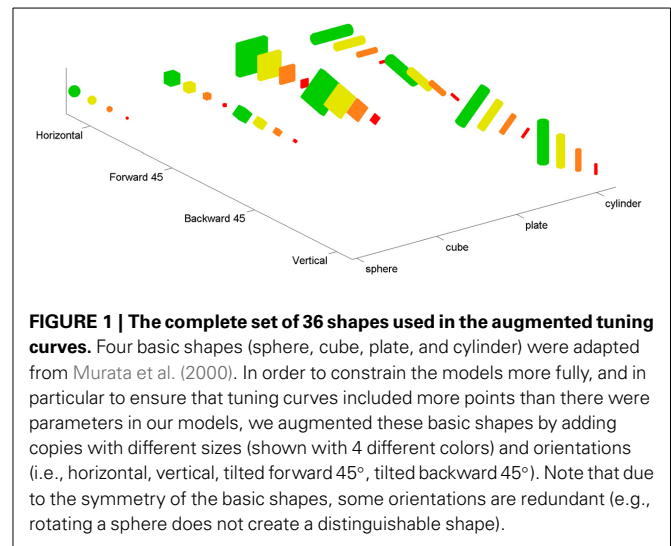
$$K_x = \frac{\partial^2 z / \partial x^2}{(1 + (\partial z / \partial x)^2)^{3/2}} \quad (4)$$

(de Vries et al., 1993). For these stimuli  $\partial z / \partial x = 0$  at the center, and so  $K_x = \partial^2 z / \partial x^2$ .

### 2.1.2. AIP Tuning

Following Sakata et al. (1998) and Murata et al. (2000) and consistent with the role of AIP in grasping (Fagg and Arbib, 1998), we took the visual-dominant neurons in AIP to be responsive to three-dimensional shape. Available tuning curves (e.g., Murata et al., 2000) span small numbers of data points relative to the large space of shape variations that are relevant to hand pre-shaping. For this reason we fit models to various “augmented” tuning curves that matched published tuning curves for some shapes, and made assumptions about how these neurons might respond to other shapes (see Figure 2). These assumptions were based on additional data for separate AIP neurons (see below). Our augmented tuning curves spanned four of the shapes in Murata et al. (2000), specifically a sphere, cylinder, cube, and plate. Two other shapes (ring and cone) were omitted for simplicity, because they require additional superquadric shape parameters (see Section 2.2). The augmented tuning curves spanned four sizes and four orientations for each of the four shapes. Due to symmetries in the shapes, there were a total of 36 points in these tuning curves (see Figure 1). Four of these points corresponded to AIP data, and the rest (the augmented points) were extrapolated from the data.

We based the augmented points on additional data from other AIP neurons, including aggregate data. Murata et al. (2000) provide shape-tuning curves for six different object-type visual-dominant AIP neurons. We tested different augmented versions of these curves with various combinations of size and orientation

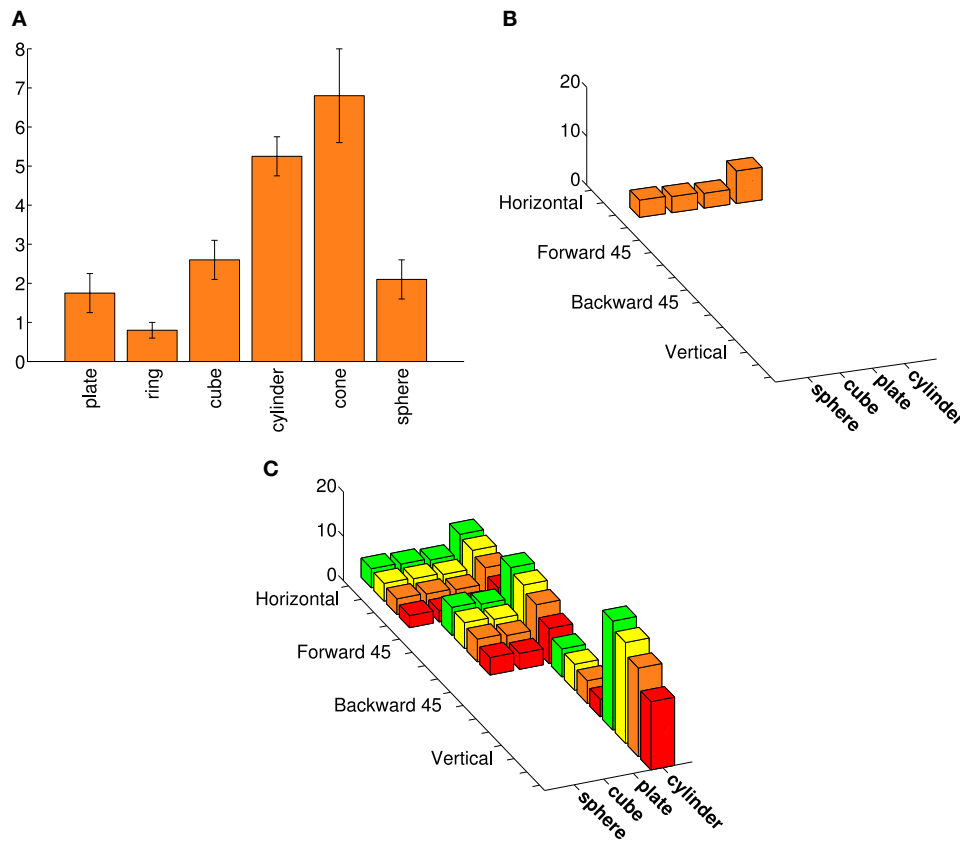


tuning (see Figure 2). Murata et al. (2000) reported (without plotting shape tuning for these neurons) that most object-type neurons were orientation selective, and that 16/26 were size-selective. Therefore, we created two augmented tuning curves for each of the six shape-tuning curves. Both were orientation-selective; one was size-selective and the other was size-invariant. For the size-selective tuning curves we assumed that spike rate increased monotonically with size (consistent with Murata et al., 2000, Figure 19; note that preference for intermediate sizes was reported only for motor-dominant neurons). We assumed that orientation tuning was roughly Gaussian and fairly narrow (consistent with Murata et al., 2000, Figure 18). Some AIP neurons are orientation selective with only mild selectivity across various elongated shapes (Sakata et al., 1998). Therefore, we created a final augmented tuning curve that was orientation selective but responded equally to cylinders and plates. Figure 1 shows an example of an augmented tuning curve and its relationship to the data. This procedure made the tuning curve optimization more challenging. This was important because even our simple cosine-tuned neuron models had more parameters than the number of points in the published tuning curves (see Section 3). It also allowed us to make use of additional AIP data.

### 2.2. SUPERQUADRICS

We modeled AIP shape tuning both on the parameters of the superquadric family of shapes, and on an Isomap dimension reduction of depth features. The superquadric family is a continuum that includes cuboids, ellipsoids, spheres, octahedra, and cylinders as examples. Superquadrics are often used to approximate observed shapes as an intermediate step in robotic grasp control (Ikeuchi and Hebert, 1996; Biegelbauer and Vincze, 2007; Goldfeder et al., 2007; Huebner et al., 2008; Duncan et al., 2013). In this context, superquadric shape parameters are typically estimated from 3D point-cloud data using iterative non-linear optimization methods (Huebner et al., 2008).

Their role in robotics suggests that superquadrics are a plausible model of AIP shape tuning. Specifically, they can be parameterized from visual information and they contain



**FIGURE 2 | An example of an augmented AIP tuning curve. (A)** Tuning curve adapted from Murata et al. (2000), Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP, 2580-2601, with permission. (See their Figure 11.) **(B)** The four points from the same tuning curve that belong to the basic superquadric family (a ring and cone are excluded from the current study). The spike rates are

plotted as 3D bars. **(C)** An augmented tuning curve that includes the points in **(B)**, as well as other rotations and scales. This augmented tuning curve is both size-tuned and orientation-tuned, as were the majority of object-type visual neurons in Murata et al. (2000). Another large minority were orientation-tuned but not size-tuned. As in **Figure 1**, the colors correspond to different sizes.

information about an object that is useful as a basis for grasp planning. One goal of the present study was to examine their physiological plausibility more closely, by fitting superquadric-tuned neuron models to AIP tuning curves. The surface of a superquadric shape is defined in  $x$ - $y$ - $z$  space as

$$\left(\frac{x}{A_1}\right)^{1/\epsilon_1} + \left(\frac{y}{A_2}\right)^{1/\epsilon_2} + \left(\frac{z}{A_3}\right)^{1/\epsilon_3} = 0,$$

where  $A > 0$  are scale parameters and  $\epsilon > 0$  are curvature parameters. Values of  $\epsilon$  close to zero correspond to squared corners, while values close to one correspond to rounded corners. For example a sphere has  $A_1 = A_2 = A_3$  and  $\epsilon_1 = \epsilon_2 = \epsilon_3 = 1$ . We also used another parameter,  $\theta$ , that described the orientation of the superquadric.  $\theta$  was composed of three angles, one per coordinate. The rotation of the superquadric is done applying the rotation matrix described in Equation 5.

$$R(\theta_1, \theta_2, \theta_3) = \begin{bmatrix} \cos(\theta_2) \cdot \cos(\theta_3) & \cos(\theta_1) \cdot \sin(\theta_3) + \sin(\theta_1) \cdot \sin(\theta_2) \cdot \cos(\theta_3) & \sin(\theta_1) \cdot \sin(\theta_3) - \cos(\theta_1) \cdot \sin(\theta_2) \cdot \cos(\theta_3) \\ -\cos(\theta_2) \cdot \sin(\theta_3) & \cos(\theta_1) \cdot \cos(\theta_3) - \sin(\theta_1) \cdot \sin(\theta_2) \cdot \sin(\theta_3) & \sin(\theta_1) \cdot \cos(\theta_3) + \cos(\theta_1) \cdot \sin(\theta_2) \cdot \sin(\theta_3) \\ \sin(\theta_2) & -\sin(\theta_1) \cdot \cos(\theta_2) & \cos(\theta_1) \cdot \cos(\theta_2) \end{bmatrix} \quad (5)$$

We generated a database of 40,000 shapes that included spheres, cylinders, plates, and cubes as well as variations on these shapes with different scales in each dimension, and rotated versions of them. Our database contained roughly equal numbers of box-like, sphere-like, and cylinder-like shapes. For round edges we set  $\epsilon = 1$ . For squared edges we drew  $\epsilon$  from an exponential distribution that was shifted slightly away from zero,  $p = 10H(\epsilon - \eta) \exp(-(\epsilon - \eta)/0.1)$  with  $\eta = 0.01$ , where  $H$  is the Heaviside step function. The shift away from 0 (perfectly sharp corners) helped to avoid numerical problems. The objects had widths between 0.02 m and 0.12 m. We also allowed arbitrary rotations in three dimensions (except where symmetry made rotations redundant), so that each shape had a total of nine parameters.

This study considers only the basic superquadric family, which does not include all the shapes for which AIP responses have been reported. However, the basic family can also be extended in various ways to deal with more complex shapes. For exam-



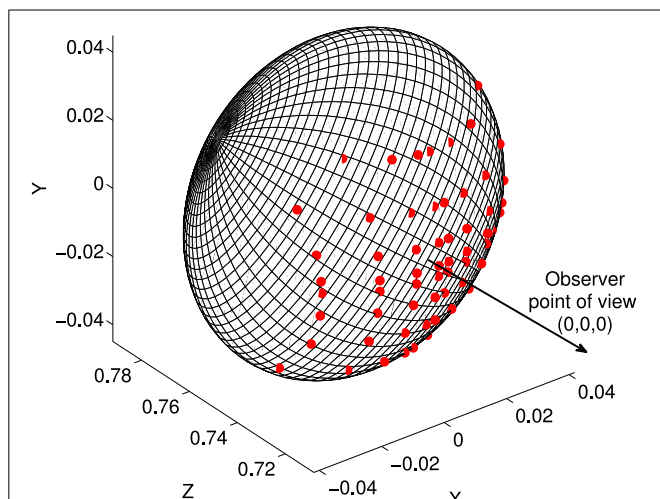
ple, hyperquadrics introduce asymmetry (Kumar et al., 1995), and trees of superquadrics can be used to approximate complex shapes with arbitrary precision (Goldfeder et al., 2007).

### 2.3. CREATION OF DEPTH MAPS

CIP receives input from V3 and V3A, which encode binocular disparity information (Anzai et al., 2011). Disparity is monotonically related to visual depth, or distance from observer to surface. As a simplified model of this input we created depth maps, i.e., grids of distances from a viewpoint to object surfaces. We created depth maps from the shapes in our superquadric database by finding intersections of the surfaces with rays at various visual angles from the view point. We used a  $16 \times 16$  grid of visual angles. Grid spacing was closer near the center than in the periphery, in order to reflect higher visual acuity near the fovea and also to ensure that a few rays intersected with the smallest shapes (specifically, distances from the center were  $a^{1.5}$ , where  $a$  were evenly-spaced points). The grid covered  $\pm 10^\circ$  of visual angle in each direction. The object centers were at a depth of 0.75 m from the viewpoint. Depth at each grid point was found as the intersection of the superquadric surface with a line from the observation point (Figure 3).

### 2.4. ISOMAP SHAPE PARAMETERS

Within the superquadric family there is typically more than one set of parameters that can describe a given shape. For example, a tall box can either be parameterized as a tall box or a wide box on its end. This is not very problematic in robotics, because an iterative search for matching parameters finds one of these solutions. However, our goal was to model a feedforward mapping from depth (V3A) to shape parameters (AIP). In order to use the superquadric parameters as the basis for an AIP tuning we therefore needed the superquadric-to-depth function to be invertible.



**FIGURE 3 | Illustration of the depth map construction process.** Each superquadric was centered at (0, 0, 0.75) relative to an observer at (0, 0, 0). Rays were traced between the observation point and a grid of points in the frontoparallel plane at  $z = 0.75$ , and intersections (red dots) were found with the superquadric surface. The depth map consisted of a grid of distances from (0, 0, 0) to these intersections.

We achieved this by restricting the ranges of angles. For example, for box-like shapes we restricted all angles to within  $\pm\pi/4$ . This resulted in a unique set of superquadric parameters for each shape. However, large discontinuities remained, in that some very similar shapes sometimes had very different parameters. For example, a tall box at an angle slightly less than  $\pi/4$  has a depth map that is very similar to a wide box at angle just greater than  $-\pi/4$  radians. Similar discontinuities seem to exist regardless of the angle convention. We anticipated that these discontinuities would impair feedforward mapping in a neural network, so we also explored an alternative low-dimensional shape parameterization.

In the alternative model, neurons were tuned to an Isomap (Tenenbaum et al., 2000) derived from depth data. Isomap is a non-linear dimension-reduction method in which samples are embedded in a lower-dimensional space in such a way that geodesic distances (i.e., distances along the shortest paths through edges between neighboring points) are maintained as well as possible. This method ensured that similar depth maps would be close together in the shape-parameter space, minimizing parameter discontinuities like those of the superquadric parameters. We constructed an Isomap of the first and second spatial derivatives of the depth maps in the horizontal and vertical directions.

We tested whether our augmented AIP tuning curves (above) were consistent with cosine tuning for these shape parameters. We also tested how well these shape parameters could be approximated by a neural network with CIP parameters as input.

### 2.5. NEURAL NETWORK MODELS OF CIP-TO-AIP MAP

In addition to fitting cosine-LIF models to neural tuning curves in CIP and AIP, we also developed feedforward networks to map from CIP variables to AIP variables. Our general approach was to decode shape parameters from the spike rates of CIP models.

We experimented with several different networks including neural engineering framework networks (Eliasmith and Anderson, 2003; Eliasmith et al., 2012), multilayer perceptrons trained with the back-propagation algorithm (Haykin, 1999) and convolutional networks (LeCun et al., 1998).

In each case the output units were linear. Linear decoding of the tuning parameters was of interest because decoding weights can be multiplied with preferred directions to give synaptic weights for any cosine tuning curve over the decoded variables (Eliasmith and Anderson, 2003). Specifically, suppose we have presynaptic rates  $\mathbf{r}_{pre}$  and linearly decoded estimates  $\hat{\mathbf{p}} = \Phi \mathbf{r}_{pre}$  of shape parameters  $\mathbf{p}$ , where  $\Phi$  is a matrix of decoding weights. In this case the family of cosine tuning curves over  $\hat{\mathbf{p}}$  is

$$r_{post} = G(\tilde{\phi}^T \hat{\mathbf{p}} + b), \quad (6)$$

where  $\tilde{\phi}^T \hat{\mathbf{p}} + b$  is the driving current,  $\tilde{\phi}$  is the neuron's preferred direction,  $G$  is a physiological model of the current-spike rate relationship, and  $b$  is a bias current. Such a tuning curve can then be obtained with synaptic weights (from all *presynaptic* neurons to a single *postsynaptic* neuron)

$$\mathbf{w}^T = \tilde{\phi}^T \Phi. \quad (7)$$

This allows us to draw general conclusions about how well our various models can account for AIP tuning, and how they would relate to future data.

Equations 6 and 7 are important components of the Neural Engineering Framework (Eliasmith and Anderson, 2003; Eliasmith et al., 2012), a method of developing large-scale neural circuit models.

### 3. RESULTS

#### 3.1. CIP TUNING

**Figure 4** shows an optimal fit of a cosine-tuned LIF model to a tuning curve from Katsuyama et al. (2010). Following their convention the spike rates are shown as a function of shape index, separately for the two curvedness levels. Inspection of the tuning curve revealed that it contained an expansive non-linearity, so we included Gaussian background noise in the model (as described in Section 2). To improve the fit further, in addition to tuning variables  $X = \partial^2 z / \partial x^2$  and  $Y = \partial^2 z / \partial y^2$  we introduced new tuning variables  $\frac{1}{2}(3(X)^2 - 1)$  and  $\frac{1}{2}(3(Y)^2 - 1)$ . The rationale for their inclusion was that these are the non-linear functions for which linear reconstruction is (with reasonable assumptions) most accurate from populations of LIF neurons tuned to  $X$  and  $Y$  (Eliasmith and Anderson, 2003). However, the fit to the Katsuyama et al. (2010) data remained poor despite these measures.

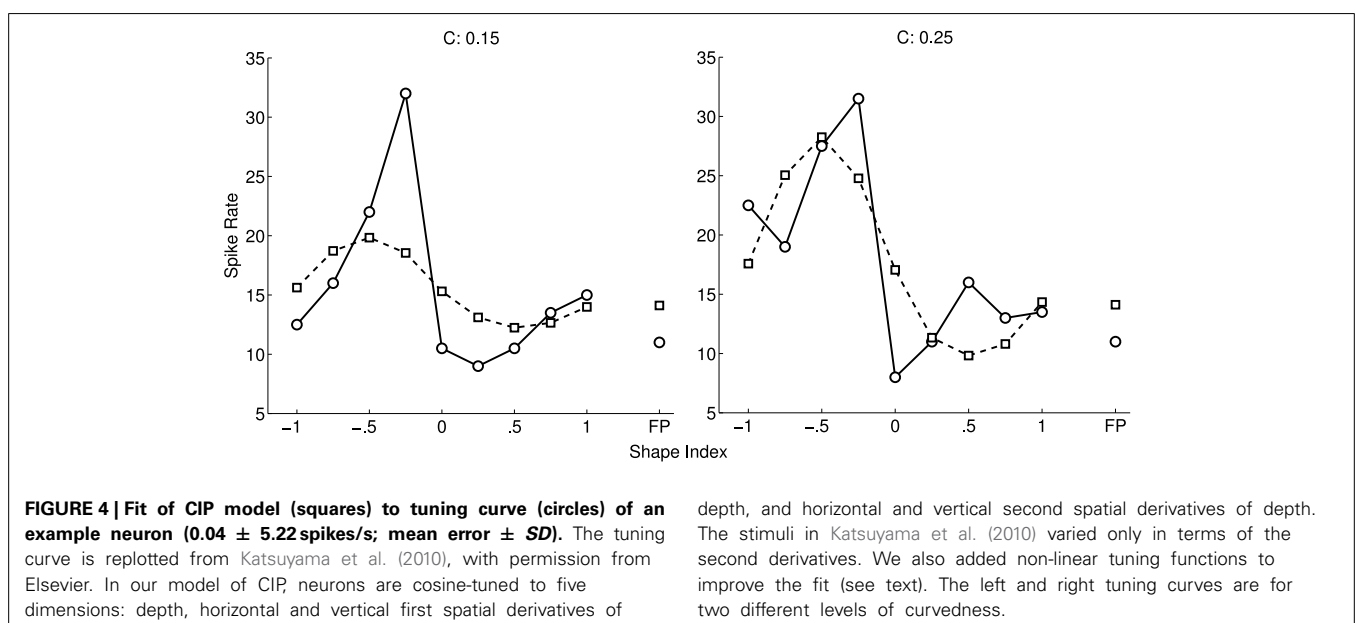
We considered whether a linear-nonlinear receptive field model with depth inputs might produce a better fit. Such models are essentially cosine tuning models with multiple input variables on a grid. However, the depth stimuli in this case (see Equation 3) consisted of linear combinations of  $x^2$  and  $y^2$ , so any receptive-field model over the depth field has an equivalent cosine tuning model over  $K_1$  and  $K_2$ . Therefore, the neuron is not cosine tuned to either depth or the curvature parameters.

**Figure 5** shows an example of a more complex non-linear neuron model that fits the data. This model is based on non-linear

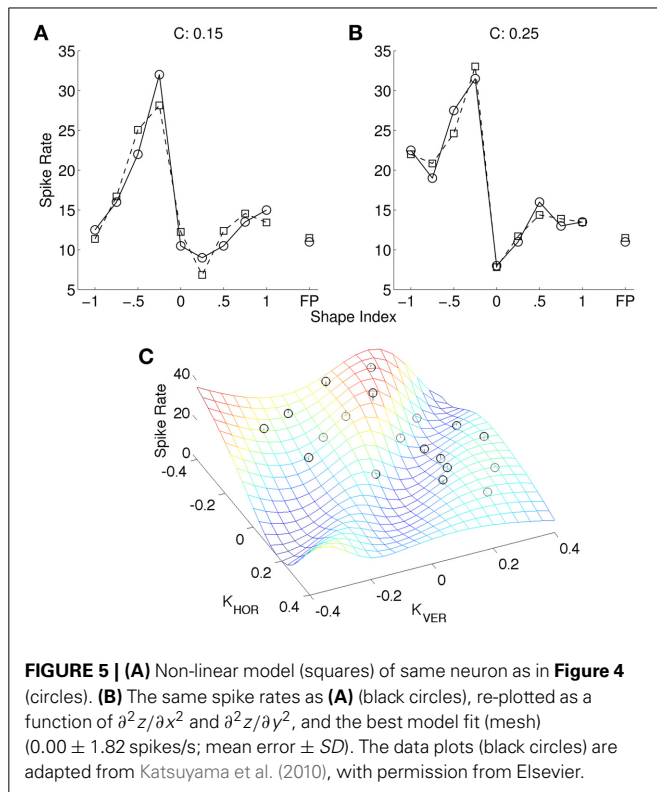
interactions between nearby inputs on the same dendrite, which suggest that pyramidal cells may function similarly to multilayer perceptrons (Polsky et al., 2004). The input to this model was a  $3 \times 3$  depth grid. The model contained 50 dendritic branches, each of which was cosine tuned to the depths. The linear kernels (analogous to preferred directions) were random. The output of each branch was a sigmoid function of the point-wise product of the depth stimulus and the linear kernel. The spike rate was a least-squares optimal weighted sum of the branch outputs. This was found using a matrix pseudoinverse that used 14 singular values. We also created another version of this model (not shown) in which the tuning curve was augmented with additional stimuli (completing the outer circle of points in **Figure 5B**) and it was assumed that the neuron would respond to these stimuli at the background spike rate. This version of the model therefore fit 26 points, and we used 20 singular values in the pseudoinverse. The fit was similar in this case.

We also constructed another alternative model of this cell that was based on a more detailed model of V3A activity. Specifically, instead of a  $3 \times 3$  depth grid, this model received input from seven non-linear functions of depth at each point. Five of these were Gaussian functions based on “tuned near,” “tuned zero,” and “tuned far” neurons (Poggio et al., 1988). Two were sigmoidal functions based on “near” and “far” tuning (Poggio et al., 1988). This model (not shown) reproduced the tuning curve somewhat less accurately than the non-linear cell model above. This was the case regardless of minor variations in the set of input tuning functions and their parameters.

**Figure 6** shows a cosine-tuning fit of data from Tsutsui et al. (2002). This tuning curve is an average over multiple cells that were tuned to depth gradients of visual stimuli. The best fitting cosine-tuning model has a notably different shape than the aggregate data. In particular, the actual spike rates are fairly constant far away from the preferred stimulus, while the model spike rates continue to decrease farther from the preferred stimulus.



Rosenberg et al. (2013) provide several additional CIP tuning curves over 49 different plane stimuli. Some of these tuning curves are clearly not consistent with cosine tuning for first derivatives of depth or disparity, e.g., with multimodal responses to surface tilt. We fit the non-linear model of **Figure 5** to seven of these tuning curves (their Figures 4, 5B). Using 20 singular values, the correlations between data and our best model fits were  $r = 0.98 \pm 0.01$  SD for the four tuning curves in their Figure 4, and  $r = 0.78 \pm 0.09$  SD for the three tuning curves in their Figure 5B. (These fits are somewhat closer than fits reported by Rosenberg et al. to Bingham functions, which is unsurprising as our model has more parameters.) Using 40 singular values, our correlations improved to  $r = 0.91 \pm 0.02$  SD for the tuning curves in their Figure 5B.



**FIGURE 5 | (A)** Non-linear model (squares) of same neuron as in **Figure 4** (circles). **(B)** The same spike rates as **(A)** (black circles), re-plotted as a function of  $\partial^2 z / \partial x^2$  and  $\partial^2 z / \partial y^2$ , and the best model fit (mesh) ( $0.00 \pm 1.82$  spikes/s; mean error  $\pm$  SD). The data plots (black circles) are adapted from Katsuyama et al. (2010), with permission from Elsevier.

In summary, the spike rates of these CIP neurons varied with the first and second spatial derivatives of depth, but not in a way that is consistent with cosine tuning to either the depth map, its first and second derivatives, or low-order polynomial functions of these derivatives. Other models, which are physiologically plausible but more complex, fit the data more closely.

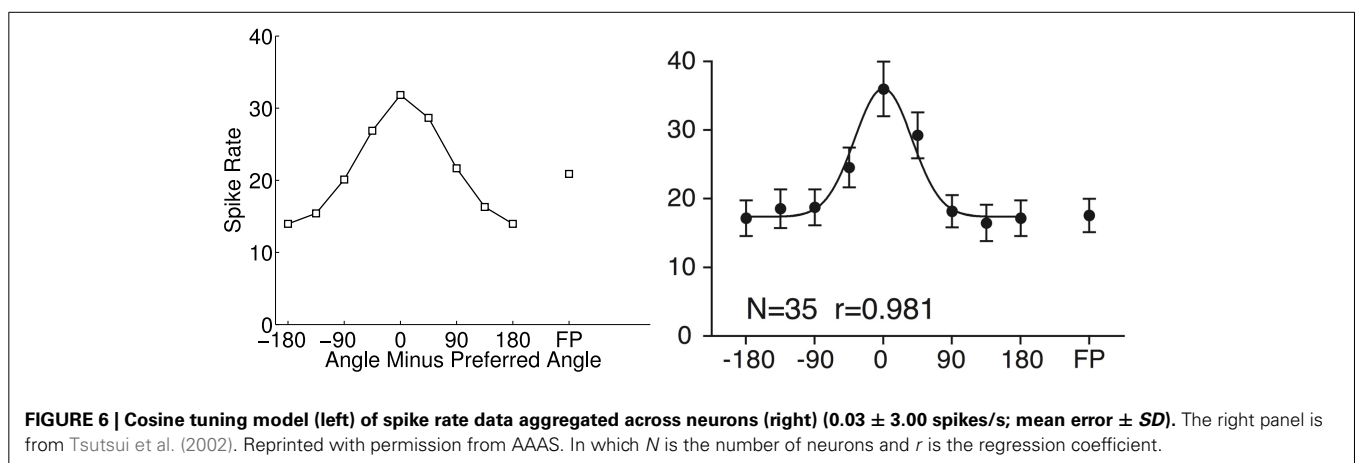
### 3.2. AIP TUNING

**Figure 7** shows an example cosine-tuning fit of an augmented tuning curve in superquadric space. This fit is based on a noise-free LIF neuron. For this dataset the shapes were rotated only in one dimension, so we avoided angle discontinuities by using a 2D direction vector in place of the angle. The optimized parameters were the 8-dimensional preferred direction vector  $\phi$ , the bias  $b$ , and the membrane time constant  $\tau_{RC}$ . Across the 36 points in the augmented tuning curve, the spike rate error (difference between augmented and model spike rates) was  $0.70 \pm 1.57$  (mean  $\pm$  SD).

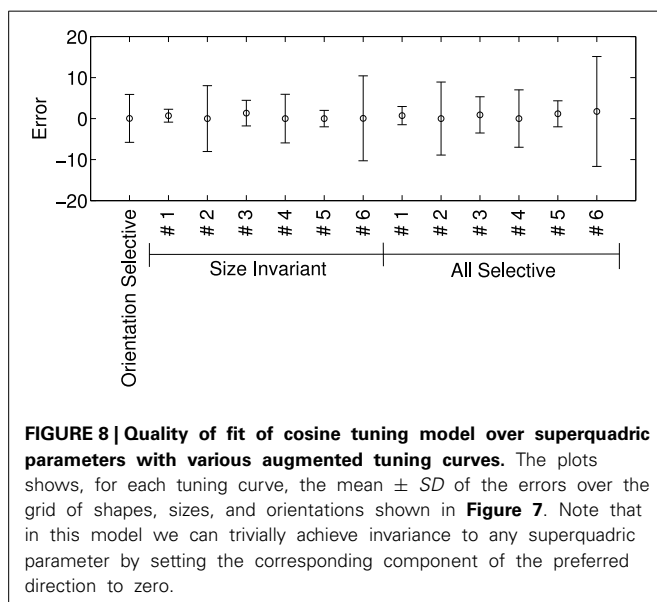
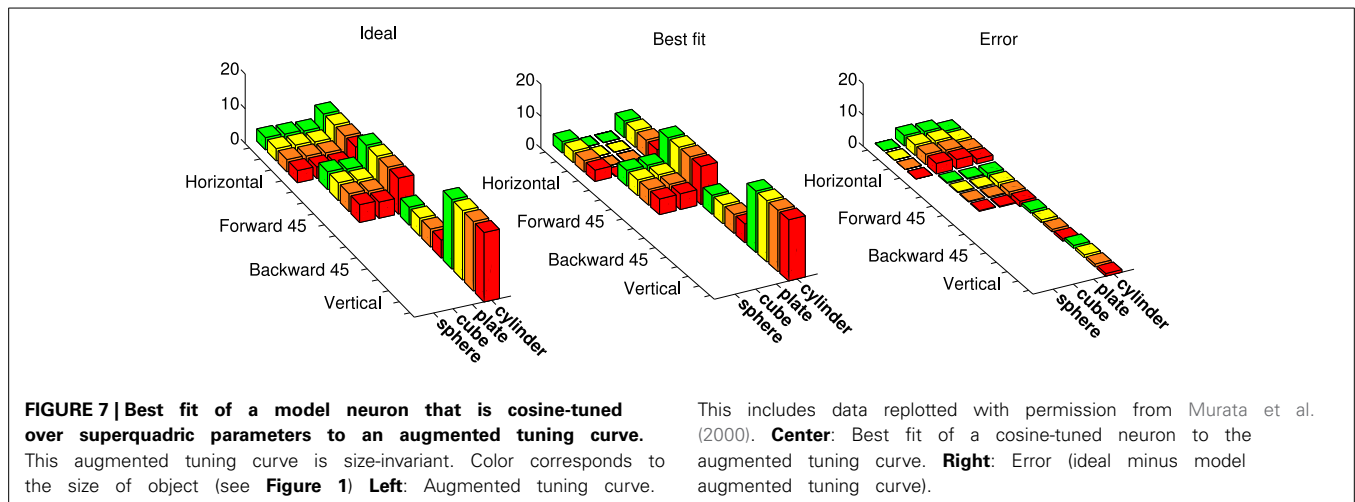
**Figure 8** shows the means and standard deviations of spike-rate errors for each of the augmented tuning curves. Good fits were obtained for some of the neurons (#1 and #3 in Murata et al., 2000, Figure 10, and the second in Figure 11, which we label #5). This was true for both size-invariant and size-selective augmented tuning curves. Neuron #1 had low spike rates for the stimuli that we studied. Neurons #3 was highly selective for cylinders, and #5 was more broadly tuned but also preferred cylinders. The worst fits were obtained for neuron #6 which responded strongly to plates and cylinders but not to cubes or spheres.

**Figure 9A** shows the means and standard deviations of spike-rate errors for each of the augmented tuning curves in an 8-dimensional Isomap space. We plot the results for the 8-dimensional Isomap in order to match the number of superquadric parameters. The cosine tuning errors ( $-0.88 \pm 10.68$  spikes/s; mean  $\pm$  SD) were larger than those in the superquadric space ( $-0.53 \pm 6.75$  spikes/s). The difference between these variances was significant according to Levene's test [ $W_{(1,910)} = 41.3$ ;  $p < 0.001$ ].

**Figure 9B** shows how the error declined with higher-dimensional Isomaps. Error variance with the 16-dimensional



**FIGURE 6 | Cosine tuning model (left) of spike rate data aggregated across neurons (right) ( $0.03 \pm 3.00$  spikes/s; mean error  $\pm$  SD). The right panel is from Tsutsui et al. (2002). Reprinted with permission from AAAS. In which  $N$  is the number of neurons and  $r$  is the regression coefficient.**



Isomap ( $-1.77 \pm 6.35$ ) was not significantly different from that of the 8-parameter superquadric [Levene's Test;  $W_{(1,910)} = 1.83$ ;  $p = 0.18$ ]. (Recalculating the variances around 0 instead of  $-1.77$  and  $-0.89$  did not make the difference significant;  $p = 0.058$ ). The cosine-tuning fits were excellent in the 32-dimensional Isomap space, with significantly lower variance [ $-0.17 \pm 1.29$  spikes/s;  $W_{(1,910)} = 316.2$ ;  $p < 0.001$ ]. This higher-dimensional shape representation is therefore consistent with the data and with the augmented tuning curves.

### 3.3. MAPPING FROM CIP TO AIP

We trained multi-layer perceptrons in order to understand whether the superquadric or Isomap models of AIP were more consistent with mapping from CIP input. Because CIP neurons are sensitive to depth and to first and second spatial derivatives of depth, we used these as inputs to the networks. Specifically

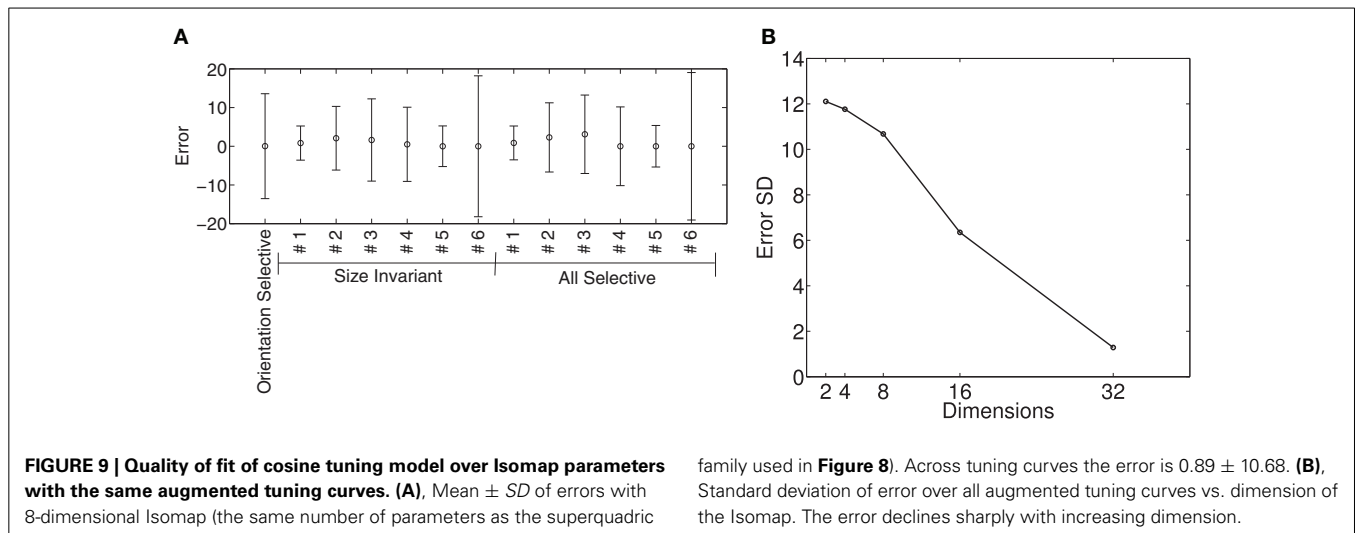
This includes data replotted with permission from Murata et al. (2000). **Center:** Best fit of a cosine-tuned neuron to the augmented tuning curve. **Right:** Error (ideal minus model augmented tuning curve).

the inputs consisted of  $16 \times 16$  depth maps, their  $16 \times 16$  horizontal and vertical derivatives, and their  $16 \times 16$  horizontal and vertical second derivatives. The derivatives were approximated by convolving with  $3 \times 3$  kernels (e.g.,  $[1 \ 1 \ 1]^T [1 \ 0 \ -1]$  and  $[1 \ 1 \ 1]^T [0.5 \ -1 \ 0.5]$ ). The total number of inputs was therefore  $16 \times 16 \times 5 = 1280$ . The hidden layers had logistic activation functions. The weights and biases were trained with the backpropagation algorithm in Matlab's Neural Network Toolbox. The output layer had a linear activation function in order to model the input to cosine-tuned neurons, as described in the Methods. A dataset of 40000 rotated superquadric objects was generated, from which depth and curvature images were derived. This dataset was divided into 28000 objects for training the network and 12000 objects to validate the results obtained in the training.

**Figure 10** shows results from networks with two hidden layers, the first with 600 units and the second with 300 units. The scatter plots show the network's output vs. the actual values of the validation dataset. In **Figure 10A** is the network's result for the superquadric shape parameter  $\epsilon_1$ . The other scatterplots in **Figures 10C,E** illustrate the network's approximation of the scale and orientation parameters  $A_1$  and  $\theta_1$ . Approximation of the other six parameters was similar (e.g., the scatterplots for  $\epsilon_2$  and  $\epsilon_3$  resemble that for  $\epsilon_1$ ). The scatterplots **Figures 10B,D,F** illustrate the network's approximation of Isomap parameters. The first, fourth, and seventh dimensions are shown as illustrative examples.

Approximation of the Isomap parameters was much more accurate than approximation of the superquadric parameters. This outcome was very consistent across a variety of networks of different sizes, with one or two hidden layers, with pre-training of hidden layers as autoencoders, etc. We also experimented with networks that contained a hidden layer of LIF neurons with random preferred directions over various local kernels, and optimal linear estimates of the shape parameters from the hidden-layer activity (Eliasmith and Anderson, 2003). The results were also similar in this case, although (as expected) more neurons were required to achieve performance like that of the more fully-optimized multilayer perceptrons.





**Figure 11** compares the distribution of the network's Isomap approximation errors with the distribution of pairwise distances between shape examples in our database. The errors were much smaller than typical distances between examples.

We also experimented with a wide variety of larger networks, including convolutional networks, using the cuda-convnet package (Krizhevsky et al., 2012). These networks did not substantially outperform the multilayer perceptron of **Figures 10, 11** (lowest mean Euclidean error 0.066 as opposed to 0.081 in **Figure 11**). We also trained some convolutional networks with only the depth map as input, and with a  $3 \times 3$  kernel in the first convolutional layer. Interestingly, some of the resulting kernels resembled the kernels that we created manually to approximate the first and second derivatives.

## 4. DISCUSSION

This study examined the neural code for three-dimensional shape in visual-dominant AIP neurons. AIP is critical for hand pre-shaping in grasping, and these neurons encode properties that are relevant to grasping including object shape, size, and orientation.

Our motivation for testing superquadric parameters as a model of AIP tuning was that superquadrics have been used in robotics, in a role that we take to be similar to the role of AIP in the primate brain. Specifically, they have been used as compact approximate representations of point clouds on which to base grasp planning. Such a representation is useful because it allows generalization from training examples to unseen examples, e.g., by interpolating between known solutions for known sets of parameters. An alternative approach in robotics is to cluster point clouds into discrete shape categories (Detry et al., 2013). We see the Isomap as an intermediate approach with some of the advantages of both superquadric fitting and clustering. The Isomap is data-driven and adapts to the statistics of the environment (like clustering), but its parameters make up a low-dimensional and continuous space (like those of superquadrics). Furthermore, unlike the superquadric representation, the Isomap representation does not have large discontinuities between very similar shapes.

We found that cosine tuning on a 32-dimensional Isomap accounted well for the tuning curves of object-selective AIP neurons. We also found that, in contrast with superquadric parameters, the Isomap parameters could be approximated fairly well by various neural networks with CIP-like input.

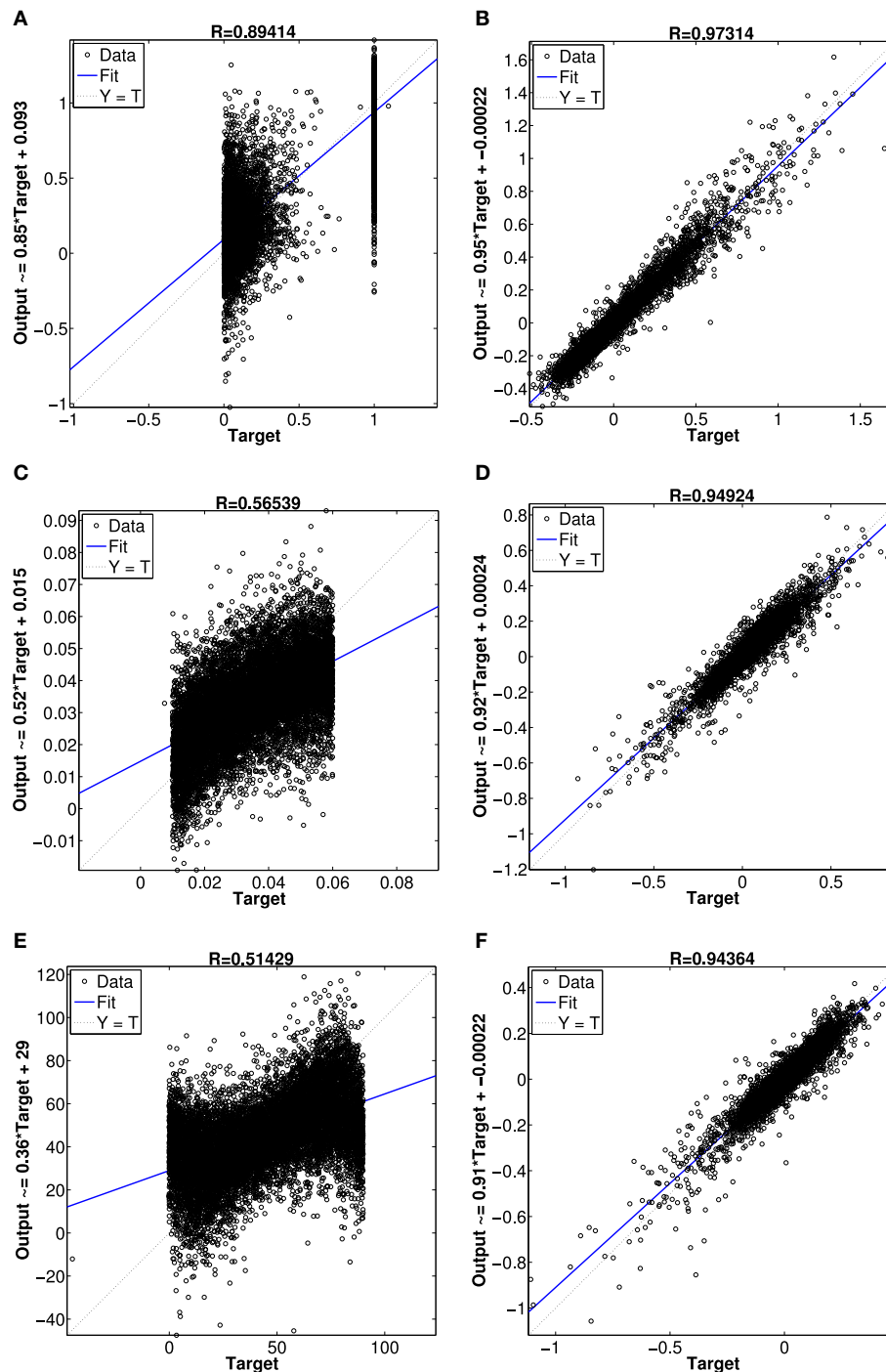
### 4.1. AUGMENTED TUNING CURVES

Available AIP data includes the responses of individual neurons to only a few different shapes, in fact fewer shapes than there are parameters in even the simplest superquadric model. To more vigorously test the different shape parameterizations as a basis for plausible neural tuning, and to incorporate additional aggregate information on shape tuning (e.g., the fact that most visual-dominant AIP neurons are orientation selective), we created “augmented” tuning curves that included both data and extrapolations of the data. It is likely that some of these augmented tuning curves were unrealistic. While the general trends in our AIP fitting results are informative (e.g., that Isomap fits improve and outperform superquadrics as dimensions increase), the details depend on our augmentation assumptions. For example, we found that the Isomap error declined more rapidly when we excluded orientation-selective/shape-invariant tuning curves from the analysis. This limitation does not affect interpretation of our other main result, i.e., that superquadrics were poorly approximated by feedforward neural networks while Isomaps were well approximated.

Future modeling would be facilitated by tuning curves with greater numbers of data points. For example, the dataset in Lehký et al. (2011) includes responses of 674 inferotemporal neurons to a common set of 806 images. A relatively extensive AIP dataset was recently collected (Schaffelhofer and Scherberger, 2014), but no tuning curves from this dataset have yet been published.

### 4.2. COSINE TUNING

We were primarily interested in cosine-tuning models for several reasons, not least because cosine tuning is widespread in the brain (see many examples in Zhang and Sejnowski, 1999). Linear-nonlinear receptive field models of the early visual system

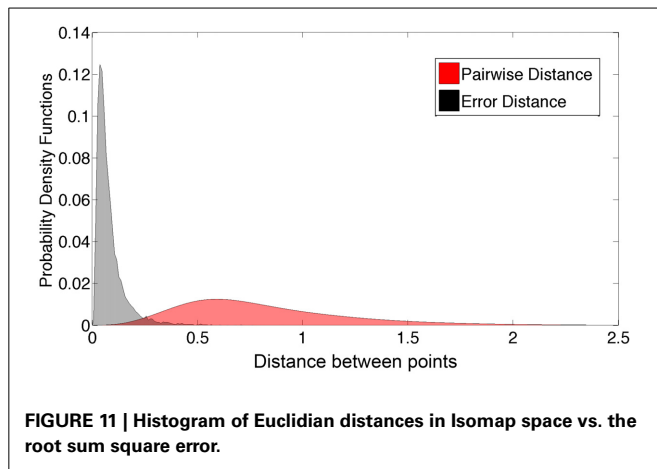


**FIGURE 10 | Regression plot comparison between neural network approximations of superquadrics parameters and Isomap parameters. (A) superquadric epsilon parameter, (B) Isomap dimension 1 parameter, (C)**

**superquadric scale parameter, (D) Isomap dimension 4 parameter, (E) superquadric rotation angle parameter and (F) Isomap dimension 7 parameter.**

are another kind of cosine tuning, with multiple tuning variables on a 2D grid. Furthermore, a practical advantage of cosine tuning models is that they require only  $n + 1$  tuning parameters for  $n$  stimulus variables (in contrast a full  $n$ -dimensional Gaussian tuning curve has  $n + n^2$  parameters). This is important because

published tuning curves in CIP and AIP consist of relatively few points, so models with large numbers of parameters may be underconstrained. Cosine tuning is also physiologically realistic in that it can arise from linear synaptic integration. For example, if a matrix  $W$  of synaptic weights has  $n$  large singular values, then



the post-synaptic neurons are tuned to a  $n$ -dimensional space (if  $W = U\Sigma V^T$  then the preferred directions are in the first  $n$  columns of  $U$ ). Cosine tuning curves are also optimal for linear decoding (Salinas and Abbott, 1994). There are also many neurons that do not appear to be cosine tuned, for example speed-tuned neurons in the middle temporal area (Nover et al., 2005). However, where applicable, cosine tuning models provide rich insight into neural activity. We therefore attempted to fit such models to the data where possible. Many AIP tuning curves over similar stimuli with different curvatures vary smoothly and monotonically (Srivastava et al., 2009), consistent with cosine tuning.

Cosine tuning to modest numbers of Isomap parameters (relative to the 256-element depth maps on which they were based) accounted for the AIP data and for our augmented AIP tuning curves.

In contrast, we concluded that the CIP neurons we modeled were not cosine tuned to the stimulus variables with which they have been examined. CIP has been proposed to encode first and second derivatives of depth (Orban et al., 2006). Various neurons in CIP respond to disparity gradient (Shikata et al., 1996; Sakata et al., 1998), texture gradient (Tsutsui et al., 2001), and/or perspective cues for oriented surfaces (Tsutsui et al., 2001). (Accordingly, visual-dominant AIP neurons also respond to monocular visual cues as well as disparity cues, and respond most strongly when disparity and other depth cues are congruent (Romero et al., 2013). Sakata et al. (1998) describe various neurons in CIP as axis-orientation-selective and surface-orientation-selective. The former were sensitive to the orientation of a long cylinder, consistent with two-dimensional tuning for horizontal and vertical curvature. The latter were selective for the orientation of a flat plate, consistent with two-dimensional tuning for depth gradient. Furthermore, Sakata et al. (1998) also recorded a neuron that preferred a cylinder of certain diameter which was tilted back and to the right, but did not respond strongly to a square column of similar dimensions. This suggests selectivity for both first and second derivatives within the same neuron. Katsuyama et al. (2010) recorded CIP responses to curved surfaces that varied in terms of their second derivatives. Tuning to the first and second derivatives of depth is physiologically plausible in that these

quantities are linear functions of the depth field, which is available from V3A. We therefore attempted to fit models that were cosine tuned over these variables, but we obtained poor fits.

While CIP neurons are certainly responsive to these variables (and more complex non-linear models of tuning to these variables fit the data closely) it is possible that there are other related variables that provide a more elegant account of these neurons' responses. Notably, some CIP neurons prefer intermediate cylinder diameters (Sakata et al., 1998), whereas cosine tuning for curvature would be constrained to monotonic changes with respect to curvature. Also, some of the neurons in Rosenberg et al. (2013) are clearly non-cosine-tuned for depth slope.

Some CIP tuning curves (see e.g., Figure 6) seem to be fairly similar to rectified cosine functions (Salinas and Abbott, 1994) with a negative offset, except that their baseline rates are not zero. In general, spike sorting limitations, which cannot be completely avoided in extracellular recordings (Harris et al., 2000), are a potential source of uncertainty in tuning curves. However, if misclassification rates had been substantial then multi-peaked tuning curves might have been expected, and none were reported in these studies.

#### 4.3. RELATIONSHIP TO SHAPE REPRESENTATION IN IT

Area IT has been shown to represent medial axes and surfaces of objects (Yamane et al., 2008; Hung et al., 2012). AIP has significant connections with IT areas including the lower bank of the superior temporal sulcus (STS), specifically areas TEa and TEm (Borra et al., 2008). These areas partially correspond to functional area TEs, which encodes curvature of depth (Janssen et al., 2000) similarly to CIP. However, AIP responds to depth differences much earlier than TEs (Srivastava et al., 2009). It is possible that a shape representation in IT, with some similarities to that in CIP, provides longer latency reinforcement and/or correction of shape representation in AIP.

#### 4.4. FUTURE WORK

A key direction for future work is to test how well the Isomap shape representation works for robotic grasp planning. This would provide important information about the functional plausibility of this representation. For example, if Isomap-based shape parameters cannot be used to shape a hand for effective grasping, this will strongly suggest that there are critical differences between AIP tuning parameters and Isomap parameters. On the other hand, if the Isomap representation performs well, it may suggest a new biologically-inspired approach for robotic grasping.

An apparent advantage of the Isomap approach is that it is data-driven and makes no prior assumptions about shapes. It would be informative to build Isomaps for less idealized shapes that monkeys might grasp in nature.

Other non-linear dimension-reduction methods (e.g., Yan et al., 2007) could also be compared with the Isomap in terms of fitting AIP data and providing an effective basis for grasp planning. We would expect differences relative to Isomap tuning to be subtle relative to available AIP data, but perhaps distinct advantages would appear in a grasp control system. One interesting possibility would be to emphasize features that are related to reward or performance (Bar-Gad et al., 2003).

Another important direction for future work is to extend the model to include motor-dominant AIP neurons and to F5 neurons as in e.g., Theys et al. (2012, 2013) and Raos et al. (2006).

Finally, our models produced constant spike rates in response to static inputs. A more sophisticated future model would account for response timing and dynamics (Sakaguchi et al., 2010). The Neural Engineering Framework (Eliasmith and Anderson, 2003) provides a principled approach to modeling dynamics in systems of spiking neurons.

## ACKNOWLEDGMENTS

Financial support was provided by CrossWing Inc.; NSERC and Mitacs (Canada); DAAD-NRF (South Africa); and the Spanish Ministry of Education and Consejo Social UPM (Spain). We thank Paul Calamai, Renaud Detry, and André Nel for helpful discussions.

## REFERENCES

- Adams, D. L., and Zeki, S. (2001). Functional organization of macaque v3 for stereoscopic depth. *J. Neurophysiol.* 86, 2195–2203. Available online at: <http://jn.physiology.org/content/86/5/2195>
- Anzai, A., Chowdhury, S. A., and DeAngelis, G. C. (2011). Coding of stereoscopic depth information in visual areas v3 and v3a. *J. Neurosci.* 31, 10270–10282. doi: 10.1523/JNEUROSCI.5956-10.2011
- Bar-Gad, I., Morris, G., and Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog. Neurobiol.* 71, 439–473. doi: 10.1016/j.pneurobio.2003.12.001
- Biegelbauer, G., and Vincze, M. (2007). “Efficient 3D object detection by fitting superquadrics to range image data for Robot’s object manipulation,” in *2007 IEEE International Conference on Robotics and Automation (ICRA)* (Piscataway, NJ: IEEE), 1086–1091. doi: 10.1109/ROBOT.2007.363129
- Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., et al. (2008). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb. Cortex* 18, 1094–1111. doi: 10.1093/cercor/bhm146
- Carandini, M. (2004). Amplification of trial-to-trial response variability by neurons in visual cortex. *PLoS Biol.* 2:E264. doi: 10.1371/journal.pbio.0020264
- Clower, D. M., Dum, R. P., and Strick, P. L. (2005). Basal ganglia and cerebellar inputs to ‘AIP’. *Cereb. Cortex* 15, 913–920. doi: 10.1093/cercor/bhh190
- Coleman, T. F., and Li, Y. (1994). On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Math. Prog.* 67, 189–224. doi: 10.1007/BF01582221
- de Vries, S. C., Kappers, A. M., and Koenderink, J. J. (1993). Shape from stereo: a systematic approach using quadratic surfaces. *Percept. Psychophys.* 53, 71–80. doi: 10.3758/BF03211716
- Detry, R., Ek, C. H., Madry, M., and Kragic, D. (2013). “Learning a dictionary of prototypical grasp-predicting parts from grasping experience,” in *2013 IEEE International Conference on Robotics and Automation (ICRA)* (Piscataway, NJ: IEEE), 601–608. doi: 10.1109/ICRA.2013.6630635
- Duncan, K., Sarkar, S., Alqasemi, R., and Dubey, R. (2013). “Multi-scale superquadric fitting for efficient shape and pose recovery of unknown objects,” in *2013 IEEE International Conference on Robotics and Automation (ICRA)* (Piscataway, NJ: IEEE), 4238–4243. doi: 10.1109/ICRA.2013.6631176
- Eliasmith, C., and Anderson, C. H. (2003). *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. Cambridge, MA: MIT Press.
- Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., et al. (2012). A large-scale model of the functioning brain. *Science* 338, 1202–1205. doi: 10.1126/science.1225266
- Fagg, A. H., and Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Netw.* 11, 1277–1303. doi: 10.1016/S0893-6080(98)00047-1
- Felleman, D. J., Burkhalter, A., and Van Essen, D. C. (1997). Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex. *J. Comp. Neurol.* 379, 21–47. doi: 10.1002/(SICI)1096-9861(19970303)379:1<21::AID-CNE3>3.0.CO;2-K
- Fitzgerald, P. J., Lane, J. W., Thakur, P. H., and Hsiao, S. S. (2004). Receptive field properties of the macaque second somatosensory cortex: evidence for multiple functional representations. *J. Neurosci.* 24, 11193–11204. doi: 10.1523/JNEUROSCI.3481-04.2004
- Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., and Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey: a reversible inactivation study. *Brain* 124(Pt 3), 571–586. doi: 10.1093/brain/124.3.571
- Gallese, V., Murata, A., Kaseda, M., Niki, N., and Sakata, H. (1994). Deficit of hand preshaping after muscimol injection in monkey parietal cortex. *Neuroreport* 5, 1525–1529. doi: 10.1097/00001756-199407000-00029
- Goldfeder, C., Allen, P. K., Lackner, C., and Pelossof, R. (2007). “Grasp planning via decomposition trees,” in *2007 IEEE International Conference on Robotics and Automation (ICRA)* (Piscataway, NJ: IEEE), 4679–4684. doi: 10.1109/ROBOT.2007.364200
- Gregoriou, G. G., Borra, E., Matelli, M., and Luppino, G. (2006). Architectonic organization of the inferior parietal convexity of the macaque monkey. *J. Comp. Neurol.* 496, 422–451. doi: 10.1002/cne.20933
- Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H., and Buzsáki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* 84, 401–414.
- Haykin, S. (1999). *Neural Networks and Learning Machines, 3rd Edn.* Upper Saddle River, NJ: Prentice Hall PTR.
- Huebner, K., Ruthotto, S., and Kragic, D. (2008). “Minimum volume bounding box decomposition for shape approximation in robot grasping,” in *2008 IEEE International Conference on Robotics and Automation (ICRA)* (Piscataway, NJ: IEEE), 1628–1633. doi: 10.1109/ROBOT.2008.4543434
- Hung, C.-C., Carlson, E., and Connor, C. (2012). Medial axis shape coding in macaque inferotemporal cortex. *Neuron* 74, 1099–1113. doi: 10.1016/j.neuron.2012.04.029
- Ikeuchi, K., and Hebert, M. (1996). “Task-oriented vision,” in *Exploratory Vision, Springer Series in Perception Engineering*, eds M. Landy, L. Maloney, and M. Pavel (New York, NY: Springer), 257–277. doi: 10.1007/978-1-4612-3984-0\_11
- Janssen, P., Vogels, R., and Orban, G. A. (2000). Three-dimensional shape coding in inferior temporal cortex. *Neuron* 27, 385–397. doi: 10.1016/S0896-6273(00)00045-3
- Jeannerod, M., Arbib, M., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320. doi: 10.1016/0166-2236(95)93921-J
- Katsuyama, N., Yamashita, A., Sawada, K., Naganuma, T., Sakata, H., and Taira, M. (2010). Functional and histological properties of caudal intraparietal area of macaque monkey. *Neuroscience* 167, 1–10. doi: 10.1016/j.neuroscience.2010.01.028
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems 25*, eds F. Pereira, C. Burges, L. Bottou, and K. Weinberger (Curran Associates, Inc.), 1097–1105.
- Krubitzer, L., Clarey, J., Tweedale, R., Elston, G., and Calford, M. (1995). A redefinition of somatosensory areas in the lateral sulcus of macaque monkeys. *J. Neurosci.* 15, 3821–3839.
- Kumar, S., Goldgof, D., and Bowyer, K. (1995). On recovering hyperquadrics from range data. *IEEE Trans. Patt. Anal. Mach. Intell.* 17, 1079–1083. doi: 10.1109/34.473234
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). “Gradient-based learning applied to document recognition,” in *Proceedings of the IEEE* (Piscataway, NJ: IEEE), 2278–2324. doi: 10.1109/5.726791
- Lehky, S. R., Kiani, R., Esteky, H., and Tanaka, K. (2011). Statistics of visual responses in primate inferotemporal cortex to object stimuli. *J. Neurophysiol.* 106, 1097–1117. doi: 10.1152/jn.00990.2010
- Luppino, G., Murata, A., Govoni, P., and Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp. Brain Res.* 128, 181–187. doi: 10.1007/s002210050833
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., and Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J. Neurophysiol.* 83, 2580–2601. Available online at: <http://jn.physiology.org/content/83/5/2580>
- Nakamura, H., Kuroda, T., Wakita, M., Kusunoki, M., Kato, A., Mikami, A., et al. (2001). From three-dimensional space vision to prehensile hand movements:



- the lateral intraparietal area links the area V3A and the anterior intraparietal area in macaques. *J. Neurosci.* 21, 8174–8187.
- Nover, H., Anderson, C. H., and DeAngelis, G. C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *J. Neurosci.* 25, 10049–10060. doi: 10.1523/JNEUROSCI.1661-05.2005
- Orban, G. A., Janssen, P., and Vogels, R. (2006). Extracting 3D structure from disparity. *Trends Neurosci.* 29, 466–473. doi: 10.1016/j.tins.2006.06.012
- Oztop, E., Imamizu, H., Cheng, G., and Kawato, M. (2006). A computational model of anterior intraparietal (AIP) neurons. *Neurocomputing* 69, 1354–1361. doi: 10.1016/j.neucom.2005.12.106
- Poggio, G. F., Gonzalez, F., and Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *J. Neurosci.* 8, 4531–4550.
- Polisky, A., Mel, B. W., and Schiller, J. (2004). Computational subunits in thin dendrites of pyramidal cells. *Nat. Neurosci.* 7, 621–627. doi: 10.1038/nn1253
- Prevete, R., Tessitore, G., Catanzariti, E., and Tamburrini, G. (2011). Perceiving affordances: a computational investigation of grasping affordances. *Cogn. Syst. Res.* 12, 122–133. doi: 10.1016/j.cogsys.2010.07.005
- Raos, V., Umiltà, M.-A., Murata, A., Fogassi, L., and Gallese, V. (2006). Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey. *J. Neurophysiol.* 95, 709–729. doi: 10.1152/jn.00463.2005
- Rizzolatti, G., Gentilucci, M., Camarda, R., Gallese, V., Luppino, G., Matelli, M., et al. (1990). Neurons related to reaching-grasping arm movements in the rostral part of area 6 (area 6a $\beta$ ). *Exp. Brain Res.* 82, 337–350. doi: 10.1007/BF00231253
- Romero, M. C., Van Dromme, I. C. L., and Janssen, P. (2013). The role of binocular disparity in stereoscopic images of objects in the macaque anterior intraparietal area. *PLoS ONE* 8:e55340. doi: 10.1371/journal.pone.0055340
- Rosenberg, A., Cowan, N. J., and Angelaki, D. E. (2013). The visual representation of 3d object orientation in parietal cortex. *J. Neurosci.* 33, 19352–19361. doi: 10.1523/JNEUROSCI.3174-13.2013
- Sakaguchi, Y., Ishida, F., Shimizu, T., and Murata, A. (2010). Time course of information representation of macaque AIP neurons in hand manipulation task revealed by information analysis. *J. Neurophysiol.* 104, 3625–3643. doi: 10.1152/jn.00125.2010
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., and Tanaka, Y. (1997). The TINS lecture. the parietal association cortex in depth perception and visual control of hand action. *Trends Neurosci.* 20, 350–357. doi: 10.1016/S0166-2236(97)01067-9
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y., and Tsutsui, K. (1998). Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1363–1373. doi: 10.1098/rstb.1998.0290
- Salinas, E., and Abbott, L. F. (1994). Vector reconstruction from firing rates. *J. Comput. Neurosci.* 1, 89–107. doi: 10.1007/BF00962720
- Schaffelhofer, S., and Scherberger, H. (2014). “From vision to action: a comparative population study of hand grasping areas AIP, F5, and M1,” in *Bernstein Conference 2014* (Göttingen). doi: 10.12751/nnn.bc2014.0253
- Shikata, E., Tanaka, Y., Nakamura, H., Taira, M., and Sakata, H. (1996). Selectivity of the parietal visual neurones in 3D orientation of surface of stereoscopic stimuli. *Neuroreport* 7, 2389–2394. doi: 10.1097/00001756-199610020-00022
- Solina, F., and Bajcsy, R. (1990). Recovery of parametric models from range images: the case for superquadrics with global deformations. *IEEE Trans. Patt. Anal. Mach. Intell.* 12, 131–147. doi: 10.1109/34.44401
- Srivastava, S., Orban, G. A., Mazire, P. A. D., and Janssen, P. (2009). A distinct representation of three-dimensional shape in macaque anterior intraparietal area: fast, metric, and coarse. *J. Neurosci.* 29, 10613–10626. doi: 10.1523/JNEUROSCI.6016-08.2009
- Taira, M., Tsutsui, K.-I., Jiang, M., Yara, K., and Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *J. Neurophysiol.* 83, 3140–3146. Available online at: <http://jn.physiology.org/content/83/5/3140>
- Tenenbaum, J. B., Silva, V. D., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* 290, 2319–2323. doi: 10.1126/science.290.5500.2319
- Theys, T., Pani, P., van Loon, J., Goffin, J., and Janssen, P. (2012). Selectivity for three-dimensional shape and grasping-related activity in the macaque ventral premotor cortex. *J. Neurosci.* 32, 12038–12050. doi: 10.1523/JNEUROSCI.1790-12.2012
- Theys, T., Pani, P., van Loon, J., Goffin, J., and Janssen, P. (2013). Three-dimensional shape coding in grasping circuits: a comparison between the anterior intraparietal area and ventral premotor area F5a. *J. Cogn. Neurosci.* 25, 352–364. doi: 10.1162/jocn-a-00332
- Tsao, D. Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., et al. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* 39, 555–568. doi: 10.1016/S0896-6273(03)00459-8
- Tsutsui, K., Jiang, M., Yara, K., Sakata, H., and Taira, M. (2001). Integration of perspective and disparity cues in surface-orientation-selective neurons of area CIP. *J. Neurophysiol.* 86, 2856–2867.
- Tsutsui, K.-I., Sakata, H., Naganuma, T., and Taira, M. (2002). Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298, 409–412. doi: 10.1126/science.1074128
- Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z., and Connor, C. E. (2008). A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat. Neurosci.* 11, 1352–1360. doi: 10.1038/nn.2202
- Yan, S., Xu, D., Zhang, B., Zhang, H.-J., Yang, Q., and Lin, S. (2007). Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Trans. Patt. Anal. Mach. Intell.* 29, 40–51. doi: 10.1109/TPAMI.2007.250598
- Zhang, K., and Sejnowski, T. J. (1999). A theory of geometric constraints on neural activity for natural three-dimensional movement. *J. Neurosci.* 19, 3122–3145.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 May 2014; accepted: 26 September 2014; published online: 27 October 2014.

Citation: Rezai O, Kleinhans A, Matallanas E, Selby B and Tripp BP (2014) Modeling the shape hierarchy for visually guided grasping. *Front. Comput. Neurosci.* 8:132. doi: 10.3389/fncom.2014.00132

This article was submitted to the journal *Frontiers in Computational Neuroscience*. Copyright © 2014 Rezai, Kleinhans, Matallanas, Selby and Tripp. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.